An International Journal
computers & mathematics with applications

PERGAMON Computers and Mathematics with Applications 45 (2003) 823–834
www.elsevier.com/locate/camwa

Solving Obstacle Problems with Guaranteed Accuracy

CHEON SEOUNG RYOO
Department of Mathematics
Hannam University
Taejon, 306-791, Korea

(Received and accepted September 2001)

Abstract—In this paper, we consider a numerical technique which enables us to verify the existence of solutions for some simple obstacle problems. Using the finite element approximation and constructive error estimates, we construct, on a computer, a set of solutions which satisfies the hypothesis of the Schauder fixed-point theorem for a compact map on a certain Sobolev space. We describe the numerical verification algorithm for solving a two-dimensional obstacle problems and report some numerical results.

Keywords—Numerical verification, Fixed-point theorem, Error estimates, Obstacle problems, Finite element method.

1. INTRODUCTION

Recently, several methods to the numerical proof of existence of solutions for variational inequalities have been developed [1–4]. The basic approach of this method consists of the fixed-point formulation of variational inequality and construction of the function set, on a computer, satisfying the validation condition of a certain infinite dimensional fixed-point theorem. The sections of this paper are as follows. In Section 2, we review mathematical background materials related to the obstacle problems, give a brief description of the fixed-point formulation, and consider methods of verification. In Section 3, a computer algorithm to construct the set satisfying the verification conditions is presented. In order to compute the rounding error, it is necessary to determine some constants which appear in *a priori* error estimates. In Section 4, we describe a method to numerically estimate such constants. Numerical examples are illustrated in the last section.

2. FORMULATION OF THE PROBLEM

Let us first set a few notations. In what follows, let Ω be a convex polygonal domain in $\mathbf{R}^2$ with a boundary $\partial\Omega$. For some integer k, let $H^k(\Omega)$ denote the $L^2(\Omega)$-Sobolev space of order k on Ω. We introduce the scalar product in $L^2(\Omega)$ by

$$(f, g) = \int_\Omega f(x)g(x)\,dx.$$

0898-1221/03/$ - see front matter Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX
PII: S0898-1221(03)00043-9

The norm in $H^k(\Omega)$ will be denoted by $\|\cdot\|_{H^k(\Omega)}$. The symbol $|\cdot|_{H^k(\Omega)}$ will stand for the seminorm

$$|u|_{H^k(\Omega)} = \left(\sum_{|\alpha|=k} \|D^\alpha u\|^2_{L^2(\Omega)} \right)^{1/2}, \qquad \|u\|_{H^k(\Omega)} = \left(\sum_{j=0}^{k} |u|^2_{H^j(\Omega)} \right)^{1/2}.$$

In addition, set $H_0^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega\}$, with an inner product $(\nabla u, \nabla v)$ for $u, v \in H_0^1(\Omega)$, and

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx, \qquad \text{where } \nabla u \cdot \nabla v = \frac{\partial u}{\partial x_1}\frac{\partial v}{\partial x_1} + \frac{\partial u}{\partial x_2}\frac{\partial v}{\partial x_2}.$$

Since we adopt $(\nabla u, \nabla v)$ as the inner product on $H_0^1(\Omega)$, the associated norm is defined by $\|u\|_{H_0^1(\Omega)} = \|\nabla u\|_{L^2(\Omega)}$. Next, we define

$$K = \left\{ v : v \in H_0^1(\Omega),\ v \geq 0 \text{ a.e. on } \Omega \right\}.$$

We now suppose the following assumptions of nonlinear function $f(\cdot)$.

ASSUMPTION 1. *f is a continuous map from $H_0^1(\Omega)$ to $L^2(\Omega)$.*

ASSUMPTION 2. *For each bounded subset $W \subset H_0^1(\Omega)$, $f(W)$ is also bounded in $L^2(\Omega)$.*

We consider the following obstacle problem:

$$\text{find } u \in K \text{ such that } a(u, v-u) \geq (f(u), v-u), \qquad \forall\, v \in K, \quad u \in K. \tag{2.1}$$

First, since $a(\cdot\,,\cdot)$ is a continuous bilinear form on $H_0^1(\Omega) \times H_0^1(\Omega)$, from the Riesz representation theorem, for each $u \in H_0^1(\Omega)$, a unique element $F(u) \in H_0^1(\Omega)$ exists such that $a(F(u), v) = (f(u), v)$, $\forall\, v \in H_0^1(\Omega)$. That is,

$$\exists\, F(u) \in H_0^1(\Omega) \text{ such that } -\Delta F(u) = f(u), \qquad \text{in } \Omega, \quad F(u) = 0, \quad \text{on } \partial\Omega. \tag{2.2}$$

Thus, the map $F : H_0^1(\Omega) \longrightarrow H_0^1(\Omega)$ is a compact operator by the above assumptions on f. In [4], problem (2.1) is equivalent to $u \in H_0^1(\Omega)$, such that

$$u = P_K F(u). \tag{2.3}$$

To verify the existence of the solution (2.1) in a computer, we use the fixed-point formulation (2.3) of a compact operator, $P_K F$, as above. The Schauder fixed-point theorem yields the existence of a solution u of problem (2.3) in some suitable set $U \subset H_0^1(\Omega)$, provided that

$$P_K F(U) \subset U. \tag{2.4}$$

Now, we describe a numerical verification method to verify the existence of solution (2.1). First, we determine a set V for a bounded, convex, and closed subset $U \subset H_0^1(\Omega)$ as

$$V = \left\{ v \in H_0^1(\Omega) : v = P_K F(u),\ \forall\, u \in U \right\}.$$

From the Schauder fixed-point theorem, if $V \subset U$ holds, then a solution (2.1) exists in the set U. Our aim is to find a set U which includes V. A procedure to verify $V \subset U$ using a computer is as follows: Let S_h be a finite-dimensional subspace of $H_0^1(\Omega)$ dependent on h $(0 < h < 1)$. We then define K_h, an approximate subset of K, by

$$K_h = S_h \cap K = \left\{ v_h : v_h \in S_h,\ v_h \geq 0 \text{ on } \overline{\Omega} \right\}.$$

Notice that K_h is a closed convex subset of S_h.

We now define the projection P_{K_h} from $H_0^1(\Omega)$ onto K_h. That is, $v_h = P_{K_h}(z)$, which is called a projection of z into K_h, and is defined as the solution to the following problem:

$$v_h \in K_h : a(v_h, \zeta - v_h) \geq a(z, \zeta - v_h), \qquad \forall \zeta \in K_h.$$

First, we note that, by the well-known result [5], for any $g \in L^2(\Omega)$, the problem

$$a(u, \psi - u) \geq (g, \psi - u), \qquad \forall \psi \in K, \quad u \in K,$$

has a unique solution $u \in H_0^1(\Omega) \cap H^2(\Omega)$ and the estimate

$$\|\Delta u\|_{L^2(\Omega)} \leq \|g\|_{L^2(\Omega)} \tag{2.5}$$

holds. Using (2.5) and error estimates, we make the following assumption as one of the approximation properties of K_h.

ASSUMPTION 3. *For each $u \in H_0^1(\Omega) \cap H^2(\Omega)$, a positive constant, C exists, independent of u and h, such that*

$$\|u - P_{K_h} u\|_{H_0^1(\Omega)} \leq Ch\|g\|_{L^2(\Omega)}. \tag{2.6}$$

Here, C is numerically determined in Section 4.

Now we define the dual cone of K_h by

$$K_h^* = \left\{w \in H_0^1(\Omega) : a(w, v) \leq 0, \ \forall v \in K_h\right\}.$$

Note that K_h^* is also a closed convex cone in $H_0^1(\Omega)$ with a vertex at 0, which is the only point common to K_h and K_h^*. We need some additional lemmas from Rodrigues (cf. [6]).

LEMMA 2.1. *Any $v \in H_0^1(\Omega)$ can be uniquely decomposed into the sum of two orthogonal elements. That is,*

$$v = P_{K_h} v \bigoplus (I - P_{K_h}) v = P_{K_h} v \bigoplus P_{K_h^*} v.$$

Here, $\bigoplus$ denotes the sum of two orthogonal elements in the sense of $H_0^1(\Omega)$.

In order to find some closed, bounded, convex subset $U \subset H_0^1(\Omega)$ satisfying (2.4), we introduce $V \subset H_0^1(\Omega)$, the rounding $R(V)$ and the rounding error $\mathrm{RE}(V)$. For any $u \in H_0^1(\Omega)$, we define the rounding $R(P_K F(u)) \in K_h$ as the solution of the following problem:

$$a(R(P_K F(u)), v_h - R(P_K F(u))) \geq (f(u), v_h - R(P_K F(u))), \qquad \forall v_h \in K_h.$$

For a set $V \subset H_0^1(\Omega)$, we define the rounding $R(V) \subset K_h$ as

$$R(V) = \{v_h \in K_h : v_h = R(P_K F(u)), \ u \in U\}.$$

Also, we define for $V \subset H_0^1(\Omega)$, the rounding error $\mathrm{RE}(V) \subset H_0^1(\Omega)$ as

$$\mathrm{RE}(V) = \left\{v \in K_h^* : \|v\|_{H_0^1(\Omega)} \leq Ch \sup_{u \in U} \|f(u)\|_{L^2(\Omega)}\right\}. \tag{2.7}$$

For a given V, we calculate the rounding $R(V) \subset K_h$ and the rounding error $\mathrm{RE}(V) \subset H_0^1(\Omega)$ such that $V \subset R(V) \bigoplus \mathrm{RE}(V)$ holds. Then, it is sufficient to find U which satisfies

$$R(V) \bigoplus \mathrm{RE}(V) \subset U.$$

Next, let us introduce the procedure for finding such a set U using computers. First, we describe how to obtain such a set $H_0^1(\Omega)$ on a computer. In order to find a set U satisfying the above condition, we use a simple iterative method. The simple iteration method is as follows.

(1) First, we obtain an approximate solution $u_h^{(0)} \in K_h$ to (2.1) by an appropriate method. Set $U_h^{(0)} = \{u_h^{(0)}\}$ and $\alpha_0 = 0$.

(2) Next, we define $R(V^{(i)})$ and $\mathrm{RE}(V^{(i)})$ for $i \geq 0$, where $V^{(i)}$ is the set defined as follows:
$$V^{(i)} = \left\{ v^{(i)} \in K : v^{(i)} = P_K F\left(u^{(i)}\right),\ u^{(i)} \in U^{(i)} \right\}.$$
$R(V^{(i)})$ is defined by the subset of K_h, which consists of all elements $v_h^{(i)} \in K_h$ such that
$$a\left(v_h^{(i)}, \psi - v_h^{(i)}\right) \geq \left(f\left(u^{(i)}\right), \psi - v_h^{(i)}\right), \qquad \forall\, \psi \in K_h, \tag{2.8}$$
holds for some $u^{(i)} \in U^{(i)}$. Let $\{\phi_j\}_{j=1\cdots M}$ be a basis of S_h such that $\phi_j(x) \geq 0$, $\forall\, x \in \Omega$. Note that $R(V^{(i)})$ can be enclosed by $R(V^{(i)}) \subset \sum_{j=1}^{M} A_j \phi_j$, where $A_j = [\underline{A_j}, \overline{A_j}]$ are intervals.

Next $\mathrm{RE}(V^{(i)})$ is defined by
$$\mathrm{RE}\left(V^{(i)}\right) = \left\{ v \in K_h^* : \|v\|_{H_0^1(\Omega)} \leq Ch \sup_{u^{(i)} \in U^{(i)}} \left\| f\left(u^{(i)}\right) \right\|_{L^2(\Omega)} \right\}.$$
Hence, $V^{(i)} \subset R(V^{(i)}) \bigoplus \mathrm{RE}(V^{(i)})$ holds.

(3) Check the verification condition
$$R\left(V^{(i)}\right) \bigoplus \mathrm{RE}\left(V^{(i)}\right) \subset U^{(i)}.$$
If the condition is satisfied, then $U^{(i)}$ is the desired set and a solution to (2.1) exists in $V^{(i)}$, and hence, in $U^{(i)}$.

(4) If the condition is not satisfied, we continue the simple iteration by using δ-inflation (i.e., let δ be a certain positive constant given beforehand), and take
$$\begin{aligned}
\alpha_{i+1} &= Ch \sup_{u^{(i)} \in U^{(i)}} \left\| f\left(u^{(i)}\right) \right\|_{L^2(\Omega)} + \delta, \\
[\alpha_{i+1}] &= \left\{ v \in K_h^* : \|v\|_{H_0^1(\Omega)} \leq \alpha_{i+1} \right\}, \\
U_h^{(i+1)} &= \sum_{j=1}^{M} \left[\underline{A_j} - \delta, \overline{A_j} + \delta\right] \phi_j, \\
U^{(i+1)} &= U_h^{(i+1)} + [\alpha_{i+1}],
\end{aligned}$$
and then go back to the second step. The reader may refer to [3] for details.

3. VERIFICATION PROCEDURES BY COMPUTER

In order to construct the set U satisfying the verification condition on a computer, we use an iterative technique similar to that in [2–4], etc. We propose a computer algorithm to obtain the set U which satisfies the verification condition.

Now we consider the following variational inequality: we set $g \in L^2(\Omega)$,
$$\text{find } u \in K \text{ such that } a(u, v-u) \geq (g, v-u), \qquad \forall\, v \in K. \tag{3.1}$$
Since $g \in L^2(\Omega)$, which implies that $u \in H^2(\Omega) \cap H_0^1(\Omega)$, and hence, that $-\Delta u - g \in L^2(\Omega)$, problem (3.1) is equivalent to the infinite-dimensional linear complementarity problem
$$\begin{aligned}
-\Delta u - g &\geq 0, &&\text{a.e. on } \Omega, \\
(-\Delta u - g) u &= 0, &&\text{a.e. on } \Omega, \\
u &\in K.
\end{aligned} \tag{3.2}$$
We now define the approximate problem corresponding to (3.1) as
$$a(u_h, v_h - u_h) \geq (g, v_h - u_h), \qquad \forall\, v_h \in K_h, \quad u_h \in K_h. \tag{3.3}$$

THEOREM 3.1. *Problem (3.3) is equivalent to the linear complementarity problem*

$$\begin{aligned} w - Dz &= -P, \qquad w \geq 0 \quad z \geq 0, \\ wz &= 0. \end{aligned} \tag{3.4}$$

Here, $D = (d_{ij})$, with $d_{ij} = (\nabla\phi_i, \nabla\phi_j)$ and $1 \leq i,\ j \leq M$, and $P \equiv ((g, \phi_j))$ is an M-dimensional vector.

PROOF. By choosing $v_h = u_h + \phi_i$ for $j = 1 \cdots M$, we have

$$a(u_h, \phi_i) \geq (g, \phi_i).$$

Setting $w = a(u_h, \phi_i) - (g, \phi_i) \geq 0$, we find the first equation

$$w_i = \sum_{j=1}^{M} d_{ij} z_j - P_i, \qquad \text{for } i = 1 \cdots M. \tag{3.5}$$

The second condition in (3.4) is the definition of K_h

$$z_i \geq 0, \qquad i = 1 \cdots M. \tag{3.6}$$

The last condition in (3.4) is obtained by choosing $v_h = \epsilon u_h$,

$$a(u_h, \epsilon u_h - u_h) \geq (g, \epsilon u_h - u_h)$$

gives

$$(\epsilon - 1) \sum_{i=1}^{M} z_i \left(\sum_{j=1}^{M} d_{ij} z_j - P_i \right) \geq 0,$$

and with the choices $\epsilon > 1$ and $1 > \epsilon > 0$ due to (3.5) and (3.6), we obtain

$$w_i z_i = 0.$$

It is easy to check, in turn, that if we take the coefficients z_j of function (3.4), then u_h is the solution of (3.3).

Let $\mathbf{R}^+$ denote the set of nonnegative real numbers. For $\alpha \in \mathbf{R}^+$, we set

$$[\alpha] \equiv \left\{ \phi \in K_h^* : \|\phi\|_{H_0^1(\Omega)} \leq \alpha \right\}. \tag{3.7}$$

For $i \geq 1$, in order to define $U^{(i)}$ in (4), we need some additional properties. Let A_j $(1 \leq j \leq M)$ be intervals on $\mathbf{R}^1$ and let $\sum_{j=1}^{M} A_j \phi_j$ be a linear combination of $\{\phi_j\}$ (i.e., an element of the power set 2^{S_h}) in the following sense:

$$\sum_{j=1}^{M} A_j \phi_j = \left\{ \sum_{j=1}^{M} a_j \phi_j : a_j \in A_j,\ 1 \leq j \leq M \right\}.$$

In order to calculate rounding $R(V)$ for a given set $U = \sum_{j=1}^{M} A_j \phi_j + [\alpha]$ and $g = f(U)$ in (3.3), we consider the nonlinear system

$$\begin{aligned} w - Dz &= -(f(U), \phi_j), \qquad 1 \leq j \leq M, \\ wz &= 0, \qquad w \geq 0, \quad z \geq 0. \end{aligned} \tag{3.8}$$

Equation (3.8) above is in fact a nonlinear system of equations whose right-hand side consists of intervals. Let $x := (w_1, \dots, w_M, z_1, \dots, z_M)$ and $n = 2M$. Many algorithms for solving (3.8) have been designed via an equivalent system of nonlinear equations

$$J(x) = 0, \tag{3.9}$$

where $J : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ is continuous. The equivalence means that x^* solves (3.8) if and only if x^* solves (3.9). In particular, inclusion methods for nonlinear equations by slope are studied in [7]. We briefly describe the method presented by Rump [7] below. Let $\mathbf{P}S$ denote the power set over a given set S. First, we construct the inclusion function as follows. Let $J : \mathbf{R}^n \to \mathbf{R}^n$ be a continuous function, and let

$$\mathcal{J} : D \to \mathbf{PR}^n \text{ satisfy } x \in D \Rightarrow J(x) \in \mathcal{J}(x). \tag{3.10}$$

Next, for a compact and convex $\emptyset \neq X$ with fixed $\tilde{x} \in D$, we assume a linearization of J with respect to some $\tilde{x}$ to be given by means of a set-valued matrix $S_J(\tilde{x}, X)$ (i.e., we suppose the existence of some function $S_J : D \times \mathbf{P}D \to \mathbf{PR}^{n \times n}$) with

$$\tilde{x} \in D, \quad X \in \mathbf{P}D \Rightarrow J(x) \in J(\tilde{x}) + S_J(\tilde{x}, X) \cdot (x - \tilde{x}). \tag{3.11}$$

In order to solve (3.8) with guaranteed accuracy, we use the following theorem which is given by Rump [7].

THEOREM 3.2. *Let $J : \mathbf{R}^n \to \mathbf{R}^n$ be a continuous function, $R \in \mathbf{R}^{n \times n}$, $\mathcal{J}$ and S_J be given according to (3.10) and (3.11), a compact and convex $\emptyset \neq X \subseteq D$ and $\tilde{x} \in D$. If, for X,*

$$\tilde{x} - R \cdot \mathcal{J}(\tilde{x}) + \{I - R \cdot S_J(\tilde{x}, X)\}(X - \tilde{x}) \overset{\circ}{\subset} X,$$

then an $\hat{x} \in X$ with $J(\hat{x}) = 0$ exists.

Let $x := (w_1, w_2, \dots, w_M,\ z_1, z_2, \dots, z_M)$ be an approximate solution of (3.8). Let I, resp. J, be the set of indices i, resp. j, for which w_i, resp. z_j, is approximately zero.

Solve the nonlinear system (3.8) using Theorem 3.2, and let W_i, $1 \le i \le M$, $i \notin I$ and Z_j, $1 \le j \le M$, $j \notin J$, be the computed inclusions for the solutions. Define $W_i := 0$ for $i \in I$ and $Z_j := 0$ for $j \in J$ and let $W := (W_1, W_2, \dots, W_M)$ and $Z := (Z_1, Z_2, \dots, Z_M)$. If then $\inf(W_i) \ge 0$ and $\inf(Z_j) \ge 0$ for $1 \le i \le M$, $1 \le j \le M$ then problem (3.8) has an optimal solution $x \in X = (W, Z)$. In actual computation of the solution for (3.8), first, we enclose a solution x of (3.9) as an interval vector $X = (W, Z)$ by application of Theorem 3.2, second, we check the condition that $\inf(W) \ge 0$ and $\inf(Z) \ge 0$.

Using the slope for a nonlinear system (3.8), we can compute the solution of (3.8) and evaluate the rounding $R(V)$.

We now consider the fully automatic generation of the set U satisfying the verification condition. First, we generate a sequence of sets $\{U^{(i)}\}$, $i = 1, 2, \dots$, which consists of subsets of $H_0^1(\Omega)$, in the following manner. We present an iterative procedure for generating $\{U^{(i)}\}$, $i = 1, 2, \dots$.

We use an iterative method with the initial value

$$U_h^{(0)} = \left\{u_h^{(0)} : u_h^{(0)} \in K_h\right\}, \qquad \alpha_0 = 0.$$

That is, for $i = 0$, we choose an appropriate initial value $u_h^{(0)} \in K_h$ and $\alpha_0 \in \mathbf{R}^+$ and define $U^{(0)} \subset H_0^1(\Omega)$ by $U^{(0)} = u_h^{(0)} \dotplus [\alpha_0]$. Usually, $u_h^{(0)}$ will be determined as

$$a\left(u_h^{(0)}, v_h - u_h^{(0)}\right) \ge \left(f\left(u_h^{(0)}\right), v_h - u_h^{(0)}\right), \qquad \forall\, v_h \in K_h, \quad u_h^{(0)} \in K_h, \tag{3.12}$$

which corresponds to the Galerkin approximation for (2.1). In addition, the standard selection for α_0 will be $\alpha_0 = 0$.

For $U_h^{(i)} = \sum_{j=1}^{M} A_j\phi_j$ and $\alpha \in \mathbf{R}^+$, we set $U^{(i)} = U_h^{(i)} + [\alpha_i]$, $i \geq 1$. Then, we define $U_h^{(i+1)} \subset K_h$ and $\alpha_{i+1} \in \mathbf{R}^+$ according to

$$\begin{aligned} w - Dz &= -\left(f\left(U^{(i)}\right), \phi_j\right), \qquad 1 \leq j \leq M, \\ wz &= 0, \qquad z \geq 0, \quad w \geq 0. \end{aligned} \tag{3.13}$$

$$\alpha_{i+1} = Ch\left\|f\left(U^{(i)}\right)\right\|_{L^2(\Omega)}, \tag{3.14}$$

where C is the same as in (2.7). Here, the $U_h^{(i+1)}$ is determined to be the solution set of (3.13), as described above. Next, we continue the simple iteration method in Section 2.

4. COMPUTATION OF THE CONSTANT C

In this section, we describe how to estimate C in (2.7). We consider the obstacle problem assuming Ω to be a convex polygonal domain of $\mathbf{R}^2$. Now we consider the following basic model problem associated with (2.1) concerning any $g \in L^2(\Omega)$:

$$a(u, v - u) \geq (g, v - u), \qquad \forall v \in K, \quad u \in K. \tag{4.1}$$

We shall now approximate the solution of (4.1) by means of a finite element method. We define the *approximation* S_h of $H_0^1(\Omega)$ by

$$S_h = \left\{v_h : v_h \in H_0^1(\Omega) \cap C^0\left(\overline{\Omega}\right),\ v_h|_T \in P_1,\ \forall T \in \mathcal{T}_h\right\},$$

where $v_h|_T$ denotes the restriction of v_h to T, where P_1 represents the space of polynomials in two variables of the degree ≤ 1 and $\mathcal{T}_h$ is the set of triangles of the triangulation. We define K_h, an approximation of K, by

$$K_h = S_h \cap K = \left\{v_h : v_h \in S_h,\ v_h \geq 0 \text{ on } \overline{\Omega}\right\}.$$

For an arbitrary solution $u \in H_0^1(\Omega)$ of (4.1) and its finite element approximation $u_h \in K_h$, defined as

$$a(u_h, v_h - u_h) \geq (g, v_h - u_h), \qquad \forall v_h \in K_h, \tag{4.2}$$

there exists a computable constant C independent of g such that

$$\|u_h - u\|_{H_0^1(\Omega)} \leq Ch\|g\|_{L^2(\Omega)}. \tag{4.3}$$

The smaller the constant C is, the higher the possibility verification is attained with the procedure described in Sections 2 and 3, as well as the higher accuracy. Now we describe how to estimate C in (4.3). Let $P_h : H_0^1(\Omega) \to S_h$ denote the H_0^1-projection defined by

$$(\nabla\zeta - \nabla(P_h\zeta), \nabla\nu) = 0, \qquad \forall \nu \in S_h.$$

To get the constructive error estimates for the projection P_h of the form, for any $\zeta \in H_0^1(\Omega) \cap H^2(\Omega)$,

$$\|\zeta - P_h\zeta\|_{H_0^1(\Omega)} \leq C_1 h|\zeta|_{H^2(\Omega)}, \tag{4.4}$$

we can apply the existing result (see [8]). We now have, by using the definition of P_h for arbitrary $\chi \in S_h^\perp$,

$$\begin{aligned} \|\zeta - P_h\zeta\|^2_{H_0^1(\Omega)} &= (\nabla(\zeta - P_h\zeta), \nabla(\zeta - P_h\zeta)) \\ &= (\nabla(\zeta - P_h\zeta), \nabla(\zeta - \chi)) \\ &\leq \|\nabla(\zeta - P_h\zeta)\|_{L^2(\Omega)}\|\nabla(u - \chi)\|_{L^2(\Omega)}, \end{aligned} \tag{4.5}$$

where $S_h^\perp$ represents the orthogonal complement of S_h in $H_0^1(\Omega)$.

Hence, choosing $\chi = \mathcal{I}\zeta$ using Nakao and Yamamoto's results [8], we may use the constant $C_1 = 0.494$ in (4.4). Here, $\mathcal{I}\zeta$ is the interpolant of ζ at each vertex. Similarly, the well-known Aubin-Nitsche's trick can also be applied to get the L^2 error by

$$\|\zeta - P_h\zeta\|_{L^2(\Omega)} \le C_1^2 h^2 |\zeta|_{H^2(\Omega)}. \tag{4.6}$$

Therefore, we have the following lemma. We take an element $\tilde{\zeta} = P_h\zeta$ as the interpolating polynomial of degree ≤ 1 of ζ.

LEMMA 4.1. *The following estimates hold for $\tilde{\zeta}$ defined above:*

$$\left\|\tilde{\zeta} - \zeta\right\|_{H_0^1(\Omega)} \le 0.494 h |\zeta|_{H^2(\Omega)}, \tag{4.7}$$

$$\left\|\tilde{\zeta} - \zeta\right\|_{L^2(\Omega)} \le (0.494)^2 h^2 |\zeta|_{H^2(\Omega)}. \tag{4.8}$$

Although the following lemma is almost the same as in [9], in order to keep this paper self-contained, we provide the details.

LEMMA 4.2. *The following estimates hold for $u \in H^2(\Omega) \cap H_0^1(\Omega)$:*

$$|v|_{H^2(\Omega)} \le \|\Delta v\|_{L^2(\Omega)}.$$

PROOF. For $v \in C_0^3(\Omega) \equiv C^3(\Omega) \cap \{v = 0 \text{ on } \partial\Omega\}$ and setting $x = (x_1, x_2)^\top$, using Green's theorem, we derive

$$\begin{aligned}\int_\Omega (\Delta v)^2\,dx &= -\sum_{i=1}^2 \int_\Omega \frac{\partial}{\partial x_i}(\Delta v)\frac{\partial v}{\partial x_i}\,dx + \int_{\partial\Omega} \Delta v \frac{\partial v}{\partial n}\,d\gamma \\ &= \sum_{i,j=1}^2 \int_\Omega \frac{\partial^2 v}{\partial x_i\,\partial x_j}\frac{\partial^2 v}{\partial x_i\,\partial x_j}\,dx + \int_{\partial\Omega}\left(\Delta v\frac{\partial v}{\partial n} - \sum_{i=1}^2 \frac{\partial^2 v}{\partial n\,\partial x_i}\frac{\partial v}{\partial x_i}\right)d\gamma,\end{aligned} \tag{4.9}$$

where $\frac{\partial v}{\partial n}$ is the outer normal derivative on $\partial\Omega$. For a differentiable point $q \in \partial\Omega$, y_1 is the tangential direction at q and y_2 is the normal direction of $\partial\Omega$. More rigorously, (y_1, y_2) is the local coordinate system at each point q of $\partial\Omega$. Let

$$I_v \equiv \Delta v\frac{\partial v}{\partial n} - \sum_{i=1}^2 \frac{\partial^2 v}{\partial n\,\partial x_i}\frac{\partial v}{\partial x_i}.$$

By the invariance of the following quantity with respect to rotations of the coordinate system $(x_1, x_2) \to (y_1, y_2)$

$$\Delta v\frac{\partial v}{\partial n} - \sum_{i=1}^2 \frac{\partial^2 v}{\partial n\,\partial x_i}\frac{\partial v}{\partial x_i},$$

we compute

$$I_v = \sum_{i=1}^2\left(\frac{\partial^2 v}{\partial y_i^2}\frac{\partial v}{\partial y_2} - \frac{\partial^2 v}{\partial y_2\,\partial y_i}\frac{\partial v}{\partial y_i}\right).$$

Since $\frac{\partial v}{\partial y_1} = 0$ at q, we obtain

$$I_v = \frac{\partial^2 v}{\partial y_1^2}\frac{\partial v}{\partial y_2} - \frac{\partial^2 v}{\partial y_2\,\partial y_1}\frac{\partial v}{\partial y_1} = \frac{\partial^2 v}{\partial y_1^2}\frac{\partial v}{\partial y_2}.$$

In fact, let $y_2 = \omega(y_1)$ be the equation of the piece of the $\partial\Omega$ the neighborhood of q. Differentiating the identity

$$v(y_1, \omega(y_1)) = 0 \tag{4.10}$$

twice with respect to y_1, we obtain

$$\frac{\partial v}{\partial y_1} + \frac{\partial v}{\partial y_2}\frac{\partial \omega}{\partial y_1} = 0,$$

$$\frac{\partial^2 v}{\partial y_1^2} + 2\frac{\partial^2 v}{\partial y_1 \partial y_2}\frac{\partial \omega}{\partial y_1} + \frac{\partial^2 v}{\partial y_2^2}\left(\frac{\partial \omega}{\partial y_1}\right)^2 + \frac{\partial v}{\partial y_2}\left(\frac{\partial^2 \omega}{\partial y_1^2}\right) = 0.$$

Since at q

$$\frac{\partial \omega}{\partial y_1} = 0,$$

it follows that

$$\frac{\partial^2 v}{\partial y_1^2} = -\frac{\partial v}{\partial y_2}\left(\frac{\partial^2 \omega}{\partial y_1^2}\right).$$

If the Ω is convex, it is not hard to see that $\frac{\partial^2 \omega}{\partial y_1^2} \le 0$. Hence,

$$I_v = -\left(\frac{\partial v}{\partial y_2}\right)^2 \frac{\partial^2 \omega}{\partial y_1^2} = -\left(\frac{\partial v}{\partial n}\right)^2 \frac{\partial^2 \omega}{\partial y_1^2} \ge 0.$$

Therefore, (4.9) can be written in the form

$$\int_\Omega (\Delta v)^2\, dx \ge \sum_{i,j=1}^{2} \int_\Omega \frac{\partial^2 v}{\partial x_i \partial x_j}\frac{\partial^2 v}{\partial x_i \partial x_j}\, dx = |v|_{H^2(\Omega)^2}.$$

In the case where Ω is polygonal domain, we have $\frac{\partial^2 \omega}{\partial y_1^2} = 0$ on $\partial\Omega$ so that $I_v = 0$. Thus, we obtain

$$\int_\Omega (\Delta v)^2\, dx = |v|^2_{H^2(\Omega)}.$$

The inequality asserted in Lemma 4.2 holds for $u \in C_0^3(\Omega)$. For $v \in H_0^1(\Omega) \cap H^2(\Omega)$, since the set of these v values is dense in $C_0^3(\Omega)$, the assertion holds for $v \in H_0^1(\Omega) \cap H^2(\Omega)$.

Then, using (2.5), Lemma 4.1 and Lemma 4.2, we have the following theorem, regarding the approximate error $\|u_h - u\|_{H_0^1(\Omega)}$.

THEOREM 4.3. *Let $\mathcal{T}_h$ be the uniform triangulation of Ω and let u and u_h be the respective solutions of (4.1) and (4.2). If $g \in L^2(\Omega)$, we then have*

$$\|u_h - u\|_{H_0^1(\Omega)} \le Ch\|g\|_{L^2(\Omega)},$$

where the constant C can be taken as 1.1045 in this case and h is the uniform mesh size of the triangle. Hence, we may take $C = 1.1045$ in (2.7).

PROOF. Since $u_h \in K_h \subset K$, results from (4.1) are that

$$a(u, u_h - u) \ge (g, u_h - u). \tag{4.11}$$

We deduce, by adding (4.2) and (4.11), that $\forall\, v_h \in K_h$,

$$a(u_h - u, u_h - u) \le a(v_h - u, u_h - u) + a(u, v_h - u) - (g, v_h - u). \tag{4.12}$$

From (4.12) and the inequality

$$2a(u,v) \le \|u\|^2_{H_0^1(\Omega)} + \|v\|^2_{H_0^1(\Omega)}, \qquad \forall\, u, v \in H_0^1(\Omega),$$

we deduce

$$\frac{1}{2}\|u_h - u\|^2_{H_0^1(\Omega)} \le \frac{1}{2}\|v_h - u\|^2_{H_0^1(\Omega)} + a(u, v_h - u) - (g, v_h - u), \qquad \forall\, v_h \in K_h. \tag{4.13}$$

Since $g \in L^2(\Omega)$ implies that $u \in H_0^1(\Omega) \cap H^2(\Omega)$ and $-\Delta u - g \in L^2(\Omega)$, if we put

$$\lambda = -\Delta u - g, \tag{4.14}$$

we get, by using (2.5),

$$\|\lambda\|_{L^2(\Omega)} \le \|\Delta u\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)} = 2\|g\|_{L^2(\Omega)}. \tag{4.15}$$

From (4.14), we deduce that

$$a(u, v) = (g + \lambda, v), \qquad \forall\, v \in H_0^1(\Omega), \tag{4.16}$$

and from (4.13) and (4.16),

$$\frac{1}{2}\|u_h - u\|^2_{H_0^1(\Omega)} \le \frac{1}{2}\|v_h - u\|^2_{H_0^1(\Omega)} + (\lambda, v_h - u), \qquad \forall\, v_h \in K_h.$$

Hence, from (4.15),

$$\frac{1}{2}\|u_h - u\|^2_{H_0^1(\Omega)} \le \frac{1}{2}\|v_h - u\|^2_{H_0^1(\Omega)} + 2\|g\|\|v_h - u\|_{L^2(\Omega)}, \qquad \forall\, v_h \in K_h. \tag{4.17}$$

To estimate $u_h - u$, we use (4.17), choosing a suitable v_h. First, we define the interpolation operator $\Pi_h : H_0^1(\Omega) \cap C^0(\Omega) \to S_h$ by

$$\begin{aligned} \Pi_h v \in S_h, \quad \forall\, v \in H_0^1(\Omega) \cap C^0\left(\overline{\Omega}\right), \\ \Pi_h v(p) = v(p), \qquad \forall\, p \in \Sigma_h. \end{aligned} \tag{4.18}$$

Here $\Sigma_h = \{p : p \in \overline{\Omega},\ p \text{ is a vertex of } T \in \mathcal{T}_h\}$. We also have

$$\Pi_h v \in K_h, \qquad \forall\, v \in K \cap C^0\left(\overline{\Omega}\right).$$

Replacing v_h by $\Pi_h u$ in (4.17), we then have

$$\frac{1}{2}\|u_h - u\|^2_{H_0^1(\Omega)} \le \frac{1}{2}\|\Pi_h u - u\|^2_{H_0^1(\Omega)} + 2\|g\|_{L^2(\Omega)}\|\Pi_h u - u\|_{L^2(\Omega)}. \tag{4.19}$$

The regularity property $u \in H^2(\Omega)$ and Lemma 4.2,

$$|u|_{H^2(\Omega)} = \|\Delta u\|_{L^2(\Omega)} \le \|g\|_{L^2(\Omega)},$$

by Lemma 4.1, we have the following estimation:

$$\|\Pi_h u - u\|_{H_0^1(\Omega)} \le 0.494h|u|_{H^2(\Omega)} \le 0.494h\|g\|_{L^2(\Omega)} \tag{4.20}$$

and

$$\|\Pi_h u - u\|_{L^2(\Omega)} \le (0.494)^2 h^2 |u|_{H^2(\Omega)} \le (0.494)^2 h^2 \|g\|_{L^2(\Omega)}. \tag{4.21}$$

From (4.19)–(4.21), we have

$$\|u_h - u\|_{H_0^1(\Omega)} \le 1.1045h\|g\|_{L^2(\Omega)}.$$

The problem of obtaining $L^2(\Omega)$ estimates of optimal order (i.e., $O(h^2)$) of $u_h - u$ via a generalization of the Aubin-Nitsche method has not yet been completely resolved. For incomplete results in this direction, we refer you to [10,11]. However, the case of two dimensions is an open problem.

5. EXAMPLE OF NUMERICAL VERIFICATION

We provide some numerical examples of verification in the two-dimensional case according to the procedures described in the previous section. Let $\Omega = (0,1) \times (0,1)$.

EXAMPLE 1. We consider the case $f(u) = 0.5u + (2\pi - 1)\sin \pi x \sin \pi y$. For simplicity, we only consider the uniform mesh here. We divide the domain into a small triangle with uniform mesh size h, and choose the basis of K_h as the pyramid function.

The execution conditions are as follows.

$\dim S_h = 100$.
Initial value: $u_h^{(0)}$ = the Galerkin approximation (3.12), $\alpha_0 = 0$.
Inflation parameters: $\delta = 10^{-3}$.

The results are as follows.

Iteration numbers for verification: 54.
$H_0^1(\Omega)$-error bound: 0.548631.
Maximum width of coefficient intervals in $\{A_j^{(N)}\} = 0.108622$.

EXAMPLE 2. Next, we consider the case $f(u) = 0.03u + \sin 2\pi x \sin 2\pi y$. The basis of K_h is the same as above. The conditions are as follows.

$\dim S_h = 100$.
Initial value: $u_h^{(0)}$ = the Galerkin approximation (3.12), $\alpha_0 = 0$.
The outline of $u_h^{(0)}$ is shown in Figure 1.
Inflation parameters: $\delta = 10^{-3}$.

The results are as follows.

Iteration numbers: 3.
$H_0^1(\Omega)$-error bound: 0.082547.
Maximum width of coefficient intervals in $\{A_j^{(N)}\} = 0.000077$.

REMARK. In the above calculations, we used typical computer arithmetic with double precision instead of strict interval computations (e.g., ACRITH-XSC, C-XSC, PROFIL, INTLAB, etc.). PROFIL is a portable C++ class fast interval library that supports an interval linear system solver proposed by Rump [12]. INTLAB is a Matlab toolbox supporting real and complex interval scalars, vectors, and matrices as well as sparse real and complex interval matrices, coded by Rump [13].

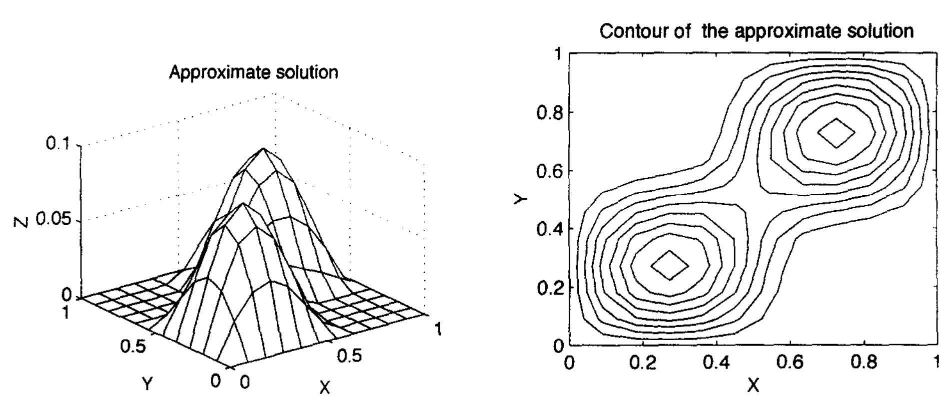

Figure 1. Approximate solution $u_h^{(0)}$.

REFERENCES

1. C.S. Ryoo and R.P. Agarwal, Numerical inclusion methods of solutions for variational inequalities, *International Journal for Numerical Methods in Engineering* **54** (11), 1545–1556, (2002).
2. M.T. Nakao, S.H. Lee and C.S. Ryoo, Numerical verification of solutions for elasto-plastic torsion problems, *Computers Math. Applic.* **39** (3/4), 195–204, (2000).
3. C.S. Ryoo, Numerical verification of solutions for a simplified Signorini problem, *Computers Math. Applic.* **40** (8/9), 1003–1013, (2000).
4. C.S. Ryoo and M.T. Nakao, Numerical verification of solutions for variational inequalities, *Numer. Math.* **81**, 305–320, (1998).
5. R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer, New York, (1984).
6. J.F. Rodrigues, *Obstacle Problems in Mathematical Physics Math. Stud. 134*, North-Holland, Amsterdam, (1987).
7. S.M. Rump, Inclusion of zeros of nowhere differentiable n-dimensional functions, *Reliable Computing* **3**, 5–16, (1997).
8. M.T. Nakao and N. Yamamoto, A guaranteed bound of the optimal constant in the error estimates for linear triangular elements, In *Perspectives on Enclosure Methods*, (Edited by V. Kulish, R. Lohner and A. Facius), pp. 265–275, Springer, (2001).
9. O.A. Ladyzhenskaya, *The Mathematical Theory of Viscous Incompressible Flow*, Gordon and Breach, New York, (1963).
10. F. Natter, Optimale L^2-Konvergenz finiten Elemente bei Variationsungleichungen, *Bonn. Math. Schr.* **89**, 1–12, (1976).
11. U. Mosco, Error estimate for some variational inequalities, *Lecture Notes in Math.* **606**, 224–236, (1977).
12. S.M. Rump, *Solving Algebraic Problems with High Accuracy. A New Approach to Scientific Computation*, Academic Press, New York, (1983).
13. S.M. Rump, *INTLAB-INTerval LABoratory*, (Edited by T. Csendes), Kluwer Academic, (1999).
14. R.P. Agarwal and J.Y. Wong, Explicit error bounds for the derivatives of piecewise-Lidstone interpolation, *J. Compt. Appl. Math.* **58**, 67–81, (1995).
15. X. Chen and T. Yamamoto, On the convergence of some quasi-Newton methods for nonlinear equations with nondifferentiable operators, *Computing* **48**, 87–94, (1992).
16. P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, (1978).
17. M.T. Nakao and N. Yamamoto, Numerical verification of solutions for nonlinear elliptic problems using an L^∞ residual method, *J. Math. Anal. Appl.* **217**, 246–262, (1998).
18. N. Yamamoto, A numerical verification method for solutions of boundary value problems with local uniqueness by Banach's fixed point theorem, *SIAM J. Numer. Anal.* **35**, 2004–2013, (1998).